

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

First Hit

End of Result Set

☐ **Generate Collection** **Print**

L9: Entry 5 of 5

File: EPAB

Jan 17, 2002

PUB-NO: DE010125406A1

DOCUMENT-IDENTIFIER: DE 10125406 A1

TITLE: Method for simultaneous access to network based electronic content using both visual and voice browsers where the voice browser calls up voice based content that can be simultaneously played or displayed with called up visual data

PUBN-DATE: January 17, 2002

INVENTOR-INFORMATION:

NAME	COUNTRY
LUCAS, BRUCE D	US
REICH, DAVID E	US

ASSIGNEE-INFORMATION:

NAME	COUNTRY
IBM	US

APPL-NO: DE10125406

APPL-DATE: May 25, 2001

PRIORITY-DATA: US60561200A (June 28, 2000)

INT-CL (IPC): G06 F 17/30; G06 F 17/20

EUR-CL (EPC): G10L015/26

ABSTRACT:

CHG DATE=20020702 STATUS=O>Method has the following steps: calling up a net based document using a visual browser, identifying a reference in the document using the voice browser, where the reference is to electronic contents in acoustic format, transfer of the reference to the voice browser, calling up the specific content using the voice browser and simultaneous visual depiction of the net-based document at the same time as the voice browser is used to depict the spoken content. An Independent claim is made for machine readable memory that can be used to run a computer program for the simultaneous access to voice bases and visual based electronic content files. An Independent claim is also made for a multi-mode browser.

h e b b g e e f c e h

e ge



⑬ **BUNDESREPUBLIK
DEUTSCHLAND**



**DEUTSCHES
PATENT- UND
MARKENAMT**

⑫ **Offenlegungsschrift**
⑩ **DE 101 25 406 A 1**

⑨ Int. Cl.⁷:
G 06 F 17/30
G 06 F 17/20

⑲ Aktenzeichen: 101 25 406.7
⑳ Anmeldetag: 25. 5. 2001
㉑ Offenlegungstag: 17. 1. 2002

DE 101 25 406 A 1

⑳ Unionspriorität:
605612 28. 06. 2000 US

㉒ Anmelder:
International Business Machines Corp., Armonk,
N.Y., US

㉓ Vertreter:
Klein, H., Rechtsanwalt, 70569 Stuttgart

㉔ Erfinder:
Lucas, Bruce D., Yorktown Heights, N.Y., US; Reich,
David E., Jupiter, Fl., US

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

Prüfungsantrag gem. § 44 PatG ist gestellt

⑤④ Verfahren und Einrichtung zum Koppeln eines Visual Browsers mit einem Voice Browser

⑤⑦ Ein Verfahren und eine Einrichtung zum gleichzeitigen Zugreifen auf netzbasierte elektronische Inhalte in einem Voice Browser und einem Visual Browser kann die Schritte des Abrufens eines netzbasierten Dokuments, das für die Darstellung im Visual Browser formatiert ist, des Identifizierens eines Bezugs auf den Voice Browser in dem abgerufenen Dokument, wobei der Bezug elektronische Inhalte spezifiziert, die für die akustische Darstellung im Voice Browser formatiert sind, und des Übertragens des Bezugs an den Voice Browser umfassen. Der Voice Browser kann die spezifizierten elektronischen Inhalte abrufen und die elektronischen Inhalte akustisch darstellen. Gleichzeitig kann der Visual Browser das netzbasierte Dokument, das für die visuelle Darstellung im Visual Browser formatiert ist, visuell darstellen. Ebenso kann das Verfahren der Erfindung die Schritte des Abrufens eines netzbasierten Dokuments, das für die akustische Darstellung im Voice Browser in dem abgerufenen Dokument, wobei der Bezug die elektronischen Inhalte spezifiziert, die für die visuelle Darstellung im Visual Browser formatiert sind, und des Übertragens des Bezugs an den Visual Browser umfassen. Der Visual Browser kann die spezifizierten elektronischen Inhalte abrufen und die spezifizierten elektronischen Inhalte visuell darstellen. Gleichzeitig kann der Voice Browser das netzbasierte Dokument, das für die akustische Darstellung im Voice Browser formatiert ist, akustisch darstellen.

DE 101 25 406 A 1

HINTERGRUND DER ERFINDUNG

5

Technisches Gebiet

[0001] Diese Erfindung betrifft das Durchsuchen von netzbasierten elektronischen Inhalten und insbesondere ein Verfahren und eine Einrichtung zum Koppeln eines Visual Browsers mit einem Voice Browser.

10

Beschreibung der zugrunde liegenden Technik

[0002] Visual Browser sind Anwendungen, die einen visuellen Zugang zu netzbasierten elektronischen Inhalten ermöglichen, die in einem Computerkommunikationsnetz zur Verfügung gestellt werden. Eine Art von Visual Browsern, der Web-Browser, dient zum Auffinden und Darstellen von netzbasierten elektronischen Inhalten, die mit der Formatierungssprache HyperText Markup Language ("HTML") formatiert sind. Zwei weit verbreitete Web-Browser sind Netscape® Navigator® und Microsoft® Internet Explorer®. Der Begriff "Visual Browser" bedeutet insbesondere, dass der Browser Grafiken, Text oder eine Kombination von Grafiken und Texten darstellen kann. Darüber hinaus können die meisten Visual Browser Multimedia-Informationen, einschließlich Klang und Video, darstellen, obwohl einige Visual Browser zur Unterstützung bestimmter Multimedia-Informationsformate Plug-Ins erfordern können.

[0003] Im Gegensatz zu typischen Visual Browsern, die in einer Desktop-Umgebung arbeiten, sind Visual Browser mit komprimierter HTML ("compressed HTML, C-HTML") zur Verarbeitung von HTML-formatierten Dokumenten in Umgebungen mit geringer Bandbreite entstanden. Genau gesagt sind C-HTML-formatierte Dokumente HTML-formatierte Dokumente, die vor der Übertragung komprimiert wurden. C-HTML-gemäße Visual Browser können C-HTML-formatierte Dokumente vor deren Darstellung dekomprimieren. Beispielhafte C-HTML-Visual Browser sind für das Betriebssystem QNX® Neutrino® von QNX Software Systems, Ltd. of Kanata, Ontario, implementiert worden.

[0004] Anders als ein Visual Browser erlaubt es ein Voice Browser einem Benutzer nicht, mit netzbasierten elektronischen Inhalten visuell zu interagieren. Vielmehr kann ein Voice Browser, der mit einer Spracherkennungseinheit und Sprachsyntheseeinheit zusammen arbeiten kann, dem Benutzer ermöglichen, mit netzbasierten elektronischen Inhalten akustisch zu interagieren. D. h., der Benutzer kann Sprachbefehle eingeben, um von einem netzbasierten elektronischen Dokument zum anderen zu navigieren. Ebenso können dem Benutzer netzbasierte elektronische Inhalte akustisch dargestellt werden, üblicherweise in Form von synthetisierter Sprache. Somit können Voice Browser Spracheingabe und interaktive Sprachausgabe für netzbasierte elektronische Inhalte und Anwendungen zur Verfügung stellen, z. B. per Telefon, PDA (Personal Digital Assistant) oder Desktop-Computer.

[0005] Es ist von Bedeutung, dass Voice Browser für die Interaktion mit netzbasierten elektronischen Inhalten, die in VoiceXML kodiert sind, konfiguriert werden können. VoiceXML ist eine Formatierungssprache für verteilte Sprachanwendungen, die auf der erweiterbaren Formatierungssprache ("Extended Markup Language, XML") basiert, ähnlich wie HTML eine Formatierungssprache für verteilte visuelle Anwendungen ist. VoiceXML ist zur Erstellung von Audio-Dialogen konzipiert, die synthetisierte Sprache, digitalisierte Töne, Erkennung von gesprochenen und Zweiton-Multifrequenz Tastatureingaben (Dual Tone Multifrequency, "DTMF"), Aufnahme gesprochener Eingabe, Telefonie und gemischt-initiierte Konversationen aufweisen. Version 1.0 der VoiceXML-Spezifikation wurde vom VoiceXML-Forum im Dokument Voice eXtensible Markup Language (VoiceXML™) version 1.0, (W3C Mai 2000) von Linda Boyer, Peter Danielsen, Jim Ferrans, Gerald Karam, David Ladd, Bruce Lucas und Kenneth Rehor veröffentlicht, das durch die Bezugnahme Bestandteil dieses Patents ist. Darüber hinaus wurde Version 1.0 der VoiceXML-Spezifikation dem World Wide Web Consortium durch das VoiceXML Forum vorgelegt und wurde von diesem als ein vorgeschlagener Industriestandard akzeptiert.

[0006] Bemerkenswerterweise wurden die Fähigkeiten von Visual Browsern nicht mit den Fähigkeiten von Voice Browsern kombiniert, damit ein Benutzer von beiden mit netzbasierten elektronischen Inhalten gleichzeitig interagieren kann. D. h., bis jetzt wurde keine Lösung zur Verfügung gestellt, die es einem Benutzer erlaubt, mit netzbasierten elektronischen Inhalten in einem Visual Browser zu interagieren, während er auch mit netzbasierten Audio-Inhalten in einem Voice Browser einen Dialog führt. Gegenwärtige Anstrengungen, einen Browser zur Verfügung zu stellen, der mit netzbasierten visuellen und akustischen Inhalten interagieren kann, sind auf die Kodierung von Sprachsynthese-Funktionalität in existierende Visual Browser beschränkt, um einen sprachsensitiven Visual Browser zu erzeugen. Außerdem wurden neue sprachbezogene Formatierungsbefehle für Visual Browser vorgeschlagen, um für einen Visual Browser Sprachfunktionalität zur Verfügung zu stellen.

[0007] Diese Lösungen erfordern immer noch, dass der, der sie implementiert, einen Satz an sprachsensitiven Funktionen zur Behandlung von netzbasierten Sprachinhalten entwickelt und diesen direkt in den Quellcode des Visual Browsers integriert. Folglich ist die Entwicklung sprachbezogener Funktionalität eng mit der Entwicklung der restlichen Funktionalität des Visual Browsers verbunden. Schließlich hindert die enge Integration des Visual Browsers und der sprachsensitiven Funktionalität den Benutzer daran, einen separaten, stabileren und effizienteren Voice Browser zu verwenden, der einen Funktionssatz hat, der für die Interaktion mit netzbasierten Sprachinhalten nützlich ist. Daher wird ein Verfahren und eine Einrichtung zum Koppeln eines Visual Browsers mit einem Voice Browser benötigt, sodass die Kombination des Visual Browsers und des Voice Browsers netzbasierte elektronische Inhalte gleichzeitig visuell und sprachlich durchsuchen kann.

65

ÜBERBLICK ÜBER DIE ERFINDUNG

[0008] Die vorliegende Erfindung ist ein Verfahren und eine Einrichtung zum gleichzeitigen Zugreifen auf netzbasierte elektronische Inhalte in einem Voice Browser und einem Visual Browser. Das Verfahren kann die Schritte des Abrufens

eines netzbasierten Dokuments, das für die Anzeige im Visual Browser formatiert ist, des Identifizierens eines Bezugs auf den Voice Browser in dem abgerufenen Dokument, wobei der Bezug elektronische Inhalte, die für die akustische Darstellung im Voice Browser formatiert sind, spezifiziert, und des Übertragens des Bezugs an den Voice Browser umfassen. Der Voice Browser kann die spezifizierten elektronischen Inhalte abrufen und die elektronischen Inhalte akustisch darstellen. Gleichzeitig kann der Visual Browser das netzbasierte Dokument, das für die visuelle Darstellung im Visual Browser formatiert ist, visuell darstellen.

[0009] In einer bevorzugten Ausführungsart der vorliegenden Erfindung kann das netzbasierte Dokument, das für die visuelle Darstellung im Visual Browser formatiert ist, ein HTML-Dokument sein und der Visual Browser kann ein Web-Browser sein. Genauso können die elektronischen Inhalte, die für die akustische Darstellung im Voice Browser formatiert sind, VoiceXML-formatierte elektronische Inhalte sein. In dem Fall, dass das netzbasierte Dokument ein HTML-Dokument ist, kann der Bezug ein Koordinationsformatierungsattribut sein.

[0010] Es ist bedeutsam, dass der Bezug ein netzbasiertes Dokument spezifizieren kann, das die elektronischen Inhalte enthält, die für die akustische Darstellung im Voice Browser formatiert sind. Folglich kann der Voice Browser das spezifizierte netzbasierte Dokument über das Computerkommunikationsnetz abrufen und die elektronischen Inhalte, die in dem spezifizierten netzbasierten Dokument enthalten sind, akustisch darstellen, während der Visual Browser das netzbasierte Dokument, das für die visuelle Darstellung formatiert ist, visuell darstellen kann.

[0011] In einer alternativen Ausführungsart kann ein Verfahren zum gleichzeitigen Zugreifen auf netzbasierte elektronische Inhalte in einem Voice Browser und einem Visual Browser die Schritte des Abrufens eines netzbasierten Dokuments, das für die akustische Darstellung im Voice Browser formatiert ist, des Identifizierens eines Bezugs auf den Visual Browser, wobei der Bezug die elektronischen Inhalte spezifiziert, die für die visuelle Darstellung im Visual Browser formatiert sind, und des Übertragens des Bezugs an den Visual Browser umfassen. Der Visual Browser kann die spezifizierten elektronischen Inhalte abrufen und die spezifizierten elektronischen Inhalte visuell darstellen. Gleichzeitig kann der Voice Browser das netzbasierte Dokument, das für die akustische Darstellung im Voice Browser formatiert ist, akustisch darstellen.

[0012] Genauer gesagt, das netzbasierte Dokument, das für akustische Darstellung im Voice Browser formatiert ist, kann ein VoiceXML-formatiertes Dokument sein. Darüber hinaus können die elektronischen Inhalte, die für die visuelle Darstellung im Visual Browser formatiert sind, HTML-formatierte elektronische Inhalte sein und der Visual Browser kann ein Web-Browser sein. Bedeutsamerweise kann der Bezug ein netzbasiertes Dokument spezifizieren, das die elektronischen Inhalte enthält, die für die visuelle Darstellung im Visual Browser formatiert sind. In diesem Fall kann der Visual Browser das spezifizierte netzbasierte Dokument über das Computerkommunikationsnetz abrufen und die elektronischen Inhalte, die in dem netzbasierten Dokument enthalten sind, visuell darstellen, während der Voice Browser das netzbasierte Dokument, das für die akustische Darstellung formatiert ist, akustisch darstellen kann.

[0013] Ein Multi-Modus-Browser für die Durchführung des Verfahrens der Erfindung kann einen Visual Browser zur visuellen Darstellung der visuellen Inhalte, die von einem Computerkommunikationsnetz abgerufen wurden, einen Voice Browser zur akustischen Darstellung der Sprachinhalte, die von dem Computerkommunikationsnetz abgerufen wurden, und eine Kopplungsschnittstelle zur Synchronisation der visuellen und akustischen Darstellung der visuellen und der Sprachinhalte im Visual bzw. Voice Browser umfassen.

[0014] Insbesondere können die visuellen Inhalte HTML-formatierte Inhalte und die Sprachinhalte VoiceXML-formatierte Inhalte sein.

[0015] Zusätzlich kann der Multi-Modus-Browser eine Steuerroutine für zeitkritische Ereignisse umfassen. Die Steuerroutine für zeitkritische Ereignisse kann zeitkritische Inhalte empfangen, in zeitkritischen Inhalten einen Inhaltstyp-Indikator identifizieren, die zeitkritischen Inhalte an den Visual Browser weiterleiten, wenn der Inhaltstyp-Indikator visuelle Inhalte anzeigt, und die zeitkritischen Inhalte an den Voice Browser weiterleiten, wenn der Inhaltstyp-Indikator Sprachinhalte anzeigt.

[0016] Ein Verfahren zum Ändern eines netzbasierten Dokuments zur Unterstützung des gleichzeitigen Zugriffs auf netzbasierte Sprach- und visuelle Inhalte in einem Voice Browser und einem Visual Browser kann die Schritte des Integrierens von visuellen Inhalten in einem netzbasierten Dokument, des Formatierens des netzbasierten Dokuments für die visuelle Darstellung in einem Visual Browser, und des Einfügens von mindestens einer Formatierungskennung in das netzbasierte Dokument umfassen. Die Formatierungskennung kann ein Koordinationsformatierungsattribut enthalten, das ein netzbasiertes Dokument, das für die akustische Darstellung in einem Voice Browser formatiert ist, spezifiziert. Als Ergebnis kann der Visual Browser bei der Wiedergabe des netzbasierten Dokuments, das für die visuelle Anzeige formatiert ist, das Koordinationsformatierungsattribut in der Formatierungskennung identifizieren und einen Bezug auf das spezifizierte netzbasierte Dokument an den Voice Browser übertragen. Folglich kann der Voice Browser das spezifizierte netzbasierte Dokument abrufen und das spezifizierte netzbasierte Dokument gleichzeitig zur visuellen Darstellung des netzbasierten Dokuments, das für die visuelle Darstellung im Visual Browser formatiert ist, akustisch darstellen. In dem bevorzugten Verfahren kann das netzbasierte Dokument, das für die visuelle Darstellung in einem Visual Browser formatiert ist, ein HTML-Dokument sein und der Visual Browser kann ein Web-Browser sein. Ebenso kann das netzbasierte Dokument, das für akustische Darstellung in einem Voice Browser formatiert ist, ein VoiceXML-formatiertes Dokument sein und der Voice Browser kann ein VoiceXML-Browser sein.

[0017] Ein Verfahren zum Ändern eines netzbasierten Dokuments zur Unterstützung des gleichzeitigen Zugriffs auf netzbasierte Sprach- und visuelle Inhalte in einem Voice Browser und einem Visual Browser kann auch die folgenden Schritte umfassen: Integrieren von Sprachinhalten in ein netzbasiertes Dokument; Formatieren des netzbasierten Dokuments für die akustische Darstellung in einem Voice Browser; und Einfügen von mindestens einer Formatierungskennung in das netzbasierte Dokument. Die Formatierungskennung kann ein Koordinationsformatierungsattribut enthalten, das ein netzbasiertes Dokument, das für die visuelle Darstellung in einem Visual Browser formatiert ist, spezifiziert. Als solches kann der Voice Browser bei der Wiedergabe des netzbasierten Dokuments, das für die akustische Darstellung formatiert ist, das Koordinationsformatierungsattribut in der Formatierungskennung identifizieren und einen Bezug auf das spezifizierte netzbasierte Dokument an den Visual Browser übertragen. Folglich kann der Visual Browser das spezifi-

zierte netzbasierte Dokument abrufen und das spezifizierte netzbasierte Dokument gleichzeitig zur akustischen Darstellung des netzbasierten Dokuments, das für die akustische Darstellung im Voice Browser formatiert ist, visuell darstellen. In einer bevorzugten Ausführungsart des Verfahrens kann das netzbasierte Dokument, das für die visuelle Darstellung im Visual Browser formatiert ist, ein HTML-formatiertes Dokument sein und der Visual Browser kann ein Web-Browser sein. Ebenso kann das netzbasierte Dokument, das für die akustische Darstellung in einem Voice Browser formatiert ist, ein VoiceXML-formatiertes Dokument sein und der Voice Browser kann ein VoiceXML-Browser sein.

KURZBESCHREIBUNG DER ZEICHNUNGEN

- [0018] In den Zeichnungen sind Ausführungsarten gezeigt, die gegenwärtig bevorzugt werden, es versteht sich jedoch, dass die Erfindung nicht auf die gezeigten exakten Anordnungen und Mittel beschränkt ist.
- [0019] Fig. 1 ist eine schematische Darstellung eines Server-Computersystems, das elektronische Inhalte zur Verfügung stellt, die für die visuelle Darstellung in einem Visual Browser auf einem Client-System formatiert sind, und eines Server-Computersystems, das elektronische Inhalte zur Verfügung stellt, die für die akustische Darstellung in einem Voice Browser auf einem Client-System formatiert sind.
- [0020] Fig. 2 ist eine hierarchische Darstellung einer Architektur, die für die Verwendung im Netz von Fig. 1 geeignet ist, in dem das Client-Computersystem einen Visual Browser und einen Voice Browser koppeln kann, sodass beide gleichzeitig auf netzbasierte elektronische Inhalte im Netz zugreifen können.
- [0021] Fig. 3 ist eine Multi-Modus-Browser-Architektur zur Implementierung des Verfahrens der Erfindung.

DETAILLIERTE BESCHREIBUNG DER ERFINDUNG

- [0022] Die vorliegende Erfindung ist ein Verfahren und eine Einrichtung zum Koppeln eines Visual Browsers mit einem Voice Browser, sodass beide gleichzeitig auf netzbasierte elektronische Inhalte in einem Computernetz auf koordinierte und sich ergänzende Weise zugreifen können. Das Verfahren und die Einrichtung der Erfindung sorgen für einen multi-modalen Aspekt, in Folge dessen die Funktionen des visuellen und sprachlichen Durchsuchens synchronisiert werden können, ohne dass ein neuer, separater Misch-Browser entwickelt werden muss und ohne dass neue Sprachbehandlungsverfahren direkt in existierende Visual Browser kodiert werden müssen. Vielmehr kann ein unabhängiger Visual Browser durch einen Bezug, der in einem netzbasierten Dokument eingebettet ist, das durch den Visual Browser abrufbar ist, an einen unabhängigen Voice Browser gekoppelt werden. Umgekehrt kann ein unabhängiger Voice Browser durch einen Bezug, der in einem netzbasierten Dokument eingebettet ist, das durch den Voice Browser abrufbar ist, an einen unabhängigen Visual Browser gekoppelt werden. In beiden Fällen werden existierende Kennungen, die durch existierende Formatierungssprachen definiert sind, verwendet, statt dass die Browser-Funktionalität geändert wird, um die Kopplung des Voice Browsers und des Visual Browsers zu ermöglichen.
- [0023] In der bevorzugten Ausführungsart kann ein netzbasierendes Dokument, das für die Anzeige in einem Visual Browser formatiert ist, von einer Quelle im Computernetz abgerufen werden. Ein Bezug auf einen Voice Browser kann im abgerufenen Dokument, in dem der Bezug elektronische Inhalte spezifiziert, die für die akustische Darstellung in dem Voice Browser formatiert sind, identifiziert werden. Anschließend kann der Bezug an den Voice Browser übertragen werden. Als Antwort auf den Empfang des übertragenen Bezugs kann der sprachgesteuerte Browser daraus die spezifizierten elektronischen Inhalte extrahieren und die elektronischen Inhalte akustisch darstellen, während der Visual Browser gleichzeitig die elektronischen Inhalte, die in dem netzbasierten Dokument enthalten sind, visuell darstellt.
- [0024] Fig. 1 zeigt eine Netzumgebung, in der das Verfahren der Erfindung durchgeführt werden kann. Diese Netzumgebung umfasst ein Computernetz 10, das einen Client-Computer 1 und die Server 20 und 30 verbindet, wobei die Server 20 und 30 mindestens einen Server für Sprachinhalte und mindestens einen Server für visuellen Inhalt umfassen. Bedeutend ist, dass die Erfindung in Hinblick auf die Anzahl der Server-Computer, die zur Versorgung des Client-Computers 1 mit netzbasierten Sprachinhalten und netzbasierten visuellen Inhalten erforderlich ist, nicht beschränkt ist. In einer Ausführungsart der vorliegenden Erfindung kann nämlich ein einzelner Server-Computer dem Client-Computer 1 sowohl Sprach- wie auch visuelle Inhalte liefern. Trotzdem wird zur leichteren Darstellung und Diskussion jeweils ein separater Server-Computer für das Bereitstellen von Sprach- und visuellen Inhalten für die anfordernden Client-Computer gezeigt.
- [0025] Das Computernetz 10 kann jedes nicht öffentlich zugängliche Netz wie z. B. ein lokales Netz (local area network, LAN) oder ein Weitverkehrsnetz (wide area network, WAN) oder vorzugsweise das Internet sein. Im Betrieb kann der Client-Computer 1 eine Verbindung mit den Server-Computern 20 und 30 herstellen, um eine Anforderung elektronischer Inhalte, die auf den Server-Computern 20 und 30 gespeichert sind, zu übertragen. Jeder der Server-Computer 20 und 30 kann auf die Anforderung durch das Bereitstellen der angeforderten elektronischen Inhalte für den Client-Computer 1 antworten. Bemerkenswerterweise können die elektronischen Inhalte für die visuelle Darstellung in einem Visual Browser oder für die akustische Darstellung in einem Voice Browser formatiert sein.
- [0026] In der bevorzugten Ausführungsart sind die elektronischen Inhalte, die für die visuelle Darstellung in einem Visual Browser formatiert sind, HTML-formatierte elektronische Inhalte, die in einem Web-Browser dargestellt werden können. Ebenso sind in der bevorzugten Ausführungsart die elektronischen Inhalte, die für die akustische Darstellung in einem Voice Browser formatiert sind, VoiceXML-formatierte elektronische Inhalte. Die Erfindung ist trotzdem nicht auf ein spezielles Verfahren der Formatierung der elektronischen Inhalte beschränkt. Vielmehr kann jedes geeignete Verfahren zur Formatierung der visuellen elektronischen Inhalte und der akustischen elektronischen Inhalte genügen. Beispiele geeigneter alternativer visueller Formatierungsverfahren können C-HTML, XML, SGML, VRML, usw. umfassen. Ebenso können Beispiele für geeignete alternative akustische Darstellungsverfahren WAV, MPEG, AIFF, etc. umfassen.
- [0027] In Fig. 1 hat der Client-Computer 1 vorzugsweise eine Zentraleinheit (central processing unit, CPU) 2, eine interne Speichereinheit 3, wie z. B. einen Arbeitsspeicher (random access memory, RAM), und eine feste Speichereinheit 4, wie etwa eine Festplatte (hard disk drive, HDD). Der Client-Computer 1 umfasst auch Netz-Schnittstellenschaltkreise

(network interface circuitry, NIC) 5 zur Kommunikationsanbindung des Client-Computers 1 an das Computerkommunikationsnetz 10. Optional kann der Client-Computer 1 des Weiteren eine Tastatur 7 und mindestens eine Benutzerschnittstellen-Anzeigeeinheit 6, wie etwa ein Videoanzeige-Terminal (video display terminal, VDT), umfassen, das daran für den Betrieb zur Interaktion mit dem Client-Computer 1 angeschlossen ist. Zum Zweck der Wiedergabe des Sprachinhalts muss der Client-Computer 1 nicht unbedingt eine Tastatur 7, eine Maus oder eine Anzeigeeinheit 6 besitzen, da diese für die Interaktion mit einem Benutzer über die Sprache nicht notwendig sind. Darüber hinaus kann der Client-Computer 1, der mit dem Benutzer akustisch einen Dialog führt, Audioschaltkreise, einen Lautsprecher und ein Mikrofon umfassen (nicht dargestellt). Des Weiteren muss der Client-Computer 1 kein Personal Computer, wie in Fig. 1 dargestellt, sein. Der Client-Computer 1 kann vielmehr jede Recheneinrichtung sein, die solche Audioschaltkreise enthält und die in der Lage ist, mit dem Computerkommunikationsnetz 10 zu kommunizieren, um die Verfahren der Erfindung auszuführen.

[0028] Wie der Client-Computer 1 haben auch die Server-Computer 20 und 30 vorzugsweise Zentraleinheiten 22 und 32, interne Speichervorrichtungen 23 und 33 und feste Speichereinheiten 24 und 34. Jeder der Server-Computer 20 und 30 umfasst auch Netz-Schnittstellenschaltkreise 25 und 35 zur Kommunikationsanbindung der Server-Computer 20 und 30 an das Computerkommunikationsnetz 10. Anders als der Client-Computer 1 müssen die Server-Computer 20 und 30 keine Tastatur oder Videoanzeigeterminale umfassen, da Benutzer mit den Server-Computern über ein Netz interagieren können. In der bevorzugten Ausführungsart kann der Server-Computer 20 dem Client-Computer 1 visuelle Inhalte für die Anzeige in einem Visual Browser, die auf dem Client-Computer 1 liegen, zur Verfügung stellen. Gleichzeitig kann der Server-Computer 30 dem Client-Computer 1 Sprachinhalte für die akustische Darstellung in einem Voice Browser auf dem Client-Computer 1 zur Verfügung stellen. Die Erfindung ist trotzdem nicht auf eine bestimmte Anzahl an zur Verfügung stehenden Server-Computern beschränkt, um elektronische Inhalte für den Client-Computer 1 bereitzustellen. In einer alternativen Ausführungsart können dem Client-Computer 1 sowohl die visuellen Inhalte als auch die Sprachinhalte durch einen einzelnen Server-Computer zur Verfügung gestellt werden.

[0029] In Fig. 2 ist eine hierarchische Darstellung einer Architektur gezeigt, die für die Verwendung im Netz von Fig. 1 geeignet ist, in der der Client-Computer einen Visual Browser 12 und einen Voice Browser 13 koppeln kann, sodass beide gleichzeitig auf netzbasierte elektronische Inhalte auf den Servern 20 und 30 im Computerkommunikationsnetz 10 zugreifen können. Wie in Fig. 2 gezeigt, kann jeder der Server-Computer 20 und 30 ein Betriebssystem (operating system, OS) 21 und 31 umfassen, auf dem die Anwendungen ausgeführt werden können. Wie nach dem Stand der Technik bekannt, kann das Betriebssystem 21 und 31 dazu verwendet werden, für eine darauf laufende Anwendung die Netzkomunikation zur Verfügung zu stellen.

[0030] Der Server-Computer 20 kann eine Server-Anwendung umfassen, die auf Anforderungen nach visuell darstellbaren elektronischen Inhalten antwortet. Genauer gesagt, die Server-Anwendung kann als Antwort auf die empfangenen Anforderungen die gewünschten elektronischen Inhalte zur Verfügung stellen. In der bevorzugten Ausführungsart ist die Server-Anwendung ein Web-Server 26, der den anfordernden Computern, die zur Kommunikation mit dem Computerkommunikationsnetz 10 verbunden sind, HTMLformatierte Inhalte zur Verfügung stellen kann. Genauer gesagt können die HTML-formatierten Inhalte statische HTML-Webseiten sein, die auf dem Server-Computer 20 gespeichert sind, oder dynamisch formatierte HTML-Webseiten, die als Antwort auf empfangene Anforderungen dynamisch erstellt werden. Alternativ dazu können die HTML-formatierten Inhalte ein komprimiertes HTML-Dokument sein.

[0031] Ebenso kann der Server-Computer 30 eine Server-Anwendung umfassen, die auf Anforderungen nach akustisch darstellbaren elektronischen Inhalten antwortet. Genauer gesagt kann die Server-Anwendung des Server-Computers 30, wie der Web-Server 26 auf dem Server-Computer 20, als Antwort auf die empfangenen Anforderungen die angeforderten elektronischen Inhalte zur Verfügung stellen. In der bevorzugten Ausführungsart ist die Server-Anwendung ein VoiceXML-Server 36, der anfordernden Computern, die zur Kommunikation mit dem Computerkommunikationsnetz 10 verbunden sind, VoiceXMLformatierte Inhalte zur Verfügung stellt. Genauer gesagt können die VoiceXML-formatierten Inhalte statische VoiceXML-Dokumente sein, die auf dem Server-Computer 30 gespeichert sind, oder dynamisch formatierte VoiceXML-Dokumente, die als Antwort auf empfangene Anforderungen dynamisch erstellt werden.

[0032] In der bevorzugten Ausführungsart kann der Client-Computer 1 ein Betriebssystem (OS) 11 umfassen, z. B. QNX Neutrino, auf dem Anwendungen geladen und ausgeführt werden können. Zwei solche bevorzugte Anwendungen können einen Visual Browser 12 und einen Voice Browser 13 einschließen. Der Visual Browser 12, z. B. ein Web-Browser, kann Benutzeranforderungen von Web-Inhalten auf dem Web-Server 26, der mit dem Computerkommunikationsnetz 10 kommunikativ verbunden ist, annehmen. Als Antwort kann der Visual Browser 12 die Anforderung an den Web-Server 26 unter Verwendung des HyperText Transfer-Protokolls ("HTTP") übertragen. Der Web-Server 26 kann die Anforderung empfangen und kann die angeforderten Web-Inhalte an den Visual Browser 12 übertragen. Ebenso kann der Voice Browser 13 Benutzeranforderungen von VoiceXML-Inhalten auf dem VoiceXML-Server 36, der mit dem Computerkommunikationsnetz 10 kommunikativ verbunden ist, annehmen. Bemerkenswert ist, dass die Benutzeranforderungen von der Spracherkennungseinheit 15 erkannt werden können. Als Antwort kann der Voice Browser 13 ebenfalls unter Verwendung von HTTP die sprachlich erkannte Anforderung an den VoiceXML-Server 36 übertragen. Der VoiceXML-Server 36 kann die Anforderung empfangen und die angeforderten VoiceXML-Inhalte an den Voice Browser 13 übertragen. Folglich können die akustisch darstellbaren elektronischen Inhalte der VoiceXML-Seite mittels der Sprachsyntheseeinheit 14 dem Benutzer akustisch dargestellt werden. Auf diese Weise kann ein Benutzer des Client-Computers 1 sowohl visuelle Inhalte mittels eines Visual Browsers 12 betrachten als auch Sprachinhalte mittels eines Voice Browsers 13 hören.

[0033] Wichtig ist, dass ein Benutzer des Client-Computers 1 gleichzeitig sowohl die visuellen als auch die Sprachinhalte, die vom Web-Server 26 bzw. VoiceXML-Server 36 zur Verfügung gestellt werden, durchsuchen kann. D. h., das Durchsuchen der visuellen Inhalte durch den Visual Browser 12 kann mit dem Durchsuchen der Sprachinhalte durch den Voice Browser 13 über eine Kopplung der visuellen Browserfunktionalität des Visual Browsers 12 mit der sprachlichen Browserfunktionalität des Voice Browsers 13 synchronisiert werden. Diese Kopplung kann erreicht werden, ohne dass das Design und die Implementierung eines Browsers für gemischte Inhalte erforderlich sind. Der Visual Browser 12 kann vielmehr an den Voice Browser 13 mittels der in den bereits existierenden Formatierungskennungen als Teil der HTML-

und VoiceXML-Spezifikation enthaltenen Koordinationsformatierungsattribute gekoppelt werden.

- [0034] Insbesondere können in der bevorzugten Ausführungsart HTMLformatierte Inhalte und VoiceXML-formatierte Inhalte Kennungen des Typs "co-target" umfassen, die zum Verweisen auf andere Browser entworfen wurden. Genauer gesagt, in einem HTML-formatierten Dokument kann eine Anker-Formatierungskennung (anchor) verwendet werden, die ein "href" enthält, um die zu ladenden Ziel-Web-Inhalte zu spezifizieren, wenn ein Benutzer den Anker auswählt. Üblicherweise wird eine URL-Adresse (uniform resource locator, URL) Ziel-Web-Inhalte in "href" spezifizieren. In der bevorzugten Ausführungsart kann ein Co-Ziel, das VoiceXML-Inhalte spezifiziert, bekannt als ein "cohref", in dem Anker zusätzlich zu den Ziel-Web-Inhalten eingebettet sein. Es folgt ein Beispiel eines durch ein "cohref" spezifizierten Co-Ziels:

```

10      <HTML>
          <HEAD>
              <TITLE>SBU Co-Browsing-Demonstration</TITLE>
15      </HEAD>
          <BODY LEFTMARGIN="0" TOPMARGIN="0">
              <A HREF="mainmenu.html"
20      cohref="mainmenu.vxml">
              <IMG SRC="splashscreen.jpg" WIDTH="314"
25      HEIGHT="200" /></A>
          </BODY>
      </HTML>

```

- [0035] Wie im obigen Beispiel gezeigt, umfasst der Hyperlink, der durch die Anker-Formatierung spezifiziert ist, einen Verweis sowohl auf eine Ziel-Webseite, die durch "mainmenu.html" spezifiziert ist, als auch auf ein Co-Ziel-VoiceXML-Dokument, das durch "mainmenu.vxml" spezifiziert ist. Wenn der Visual Browser 12 auf diesen Hyperlink stößt, kann der Visual Browser 12 vom Web-Server 26 die Web-Seite "mainmenu.html" anfordern. Darüber hinaus kann der Visual Browser 12 den Co-Ziel-Bezug identifizieren und den Bezug an den Voice Browser 13 übertragen. Der Visual Browser 12 kann den Bezug an den Voice Browser 13 unter Verwendung der Methode visit(), die in der veröffentlichten Anwendungsprogrammierschnittstelle (applications programming interface, API) des Voice Browsers 13 enthalten ist, übertragen. Als Antwort kann der Voice Browser 13 das VoiceXML-Dokument "mainmenu.vxml" vom VoiceXML-Server 36 anfordern.

- [0036] Entsprechend kann in einem VoiceXML-formatierten Dokument eine Formatierungskennung Nächste Auswahl ("choice next") verwendet werden, um die zu ladenden Ziel-VoiceXML-Inhalte zu spezifizieren, wenn ein Benutzer sprachlich eine bestimmte Auswahl in dem VoiceXML-Dokument trifft. Üblicherweise werden die Ziel-Web-Inhalte durch eine URL-Adresse ("URL") in der Kennung "choice next" spezifiziert. In der bevorzugten Ausführungsart kann ein Attribut "conext", das Web-Inhalte spezifiziert, in der Kennung "choice next" zusätzlich zu den Ziel-VoiceXML-Inhalten eingebettet sein. Es folgt ein Beispiel einer Kennung "choice next", die ein Attribut "conext" enthält:


```

<vxml version="1.0">
  <form name="Willkommen">
    <block>Willkommen bei der Co-Browsing-
Demonstration!
      <goto next="#1"/>
    </block>
  </form>
  <menu id="1">
    <prompt>Sagen Sie zur Authentifizierung
    des Benutzers "Test".</prompt>
    <grammar type="text/jsgf">
      continue
    </grammar>
    <choice next="mainmenu.vxml"
conext="mainmenu.html">
      test
    </choice>
  </menu>
</vxml>

```

[0037] Wie im obigen Beispiel gezeigt, umfasst der Hypertlink, der durch die Kennung "next" spezifiziert ist, sowohl einen Verweis auf ein Ziel-VoiceXML-Dokument, das durch "mainmenu.vxml" spezifiziert ist, als auch einen Verweis auf eine Co-Ziel-Webseite, die durch "mainmenu.html" spezifiziert ist. Wenn der Voice Browser 13 auf dieses "next" trifft, kann der Voice Browser 13 das VoiceXML-Dokument "mainmenu.vxml" vom Sprach-Server 36 anfordern. Darüber hinaus kann der Voice Browser 13 den Bezug "conext" identifizieren und den Bezug an den Visual Browser 12 übertragen. Der Voice Browser 13 kann den Bezug an den Visual Browser 12 unter Verwendung eines veröffentlichten API-Funktionsaufrufs des Visual Browsers 12 übertragen, um eine Web-Seite durch die URL der Web-Seite aufzurufen, z. B. die Funktion HTTP "Open". Der Visual Browser 12 kann wiederum die Web-Seite "mainmenu.html" vom Web-Server 26 anfordern.

[0038] Bedeutsamerweise ist die vorliegende Erfindung nicht nur auf die Koordinationsformatierungsattribute "href" und "conext" beschränkt. Es stehen vielmehr eine Anzahl an Koordinationsformatierungsattributen für die Verwendung in dem Visual Browser 12 und dem Voice Browser 13 zur Koordinierung der Aktionen des Visual Browsers 12 mit den Aktionen des Voice Browsers 13 zur Verfügung. Genauer gesagt können die Attribute in zwei Kategorien aufgeteilt werden. Erstens können Attribute Aktionen reflektieren, die an der Quelle eines Ereignisses auftreten, wie etwa das Auswählen eines Sprachmenüeintrags, das Vervollständigen eines Sprachformulars oder das Klicken auf einen visuellen Link oder eine Schaltfläche. Zweitens können Attribute Aktionen reflektieren, die am Ziel eines Ereignisses auftreten können, wie etwa das Aufsuchen eines Sprachkörpers, -menüs oder -formulars, oder das Anzeigen eines visuellen Dokuments oder eines Ankers in einem Dokument. Es ist anzumerken, dass Koordinationsformatierungsattribute im Hinblick auf den Visual Browser 12 definiert werden können, die standardmäßig nicht Teil der HTML-Spezifikation sind. Folglich können diese nicht standardmäßigen Attribute in einem anpassbaren Visual Browser, der einen entsprechend erweiterten Kennungsattributsatz besitzt, definiert werden. In der bevorzugten Ausführungsart stellt die folgende Tabelle die bevorzugten Koordinationsformatierungsattribute zur Verwendung in dem Visual Browser 12 und dem Voice Browser 13 der vorliegenden Erfindung dar:

	An der Quelle	Am Ziel
5	Voice Browser	
	<goto next="URL" conext="URL">	<menu covisit="URL">
	<choice next="URL" conext="URL">	<form covisit="URL">
10	<form action="URL" coaction="URL">	
15	Visual Browser	
		<body covisit="URL">
	<form action="URL" coaction="URL">	<form covisit="URL">
20		

[0039] Im Hinblick auf den Voice Browser 13 spezifiziert "conext" die anzuzeigende URL für den Visual Browser 12, wenn ein zugehöriges sprachlich spezifiziertes "goto" oder eine Menüauswahl durch einen Benutzer ausgewählt wird. "Coaction" spezifiziert für den Visual Browser 12 die anzuzeigende URL, wenn ein zugehöriges Sprachformular vervollständigt wurde. Schließlich spezifiziert "covisit" für den Visual Browser 12 die aufzusuchende URL, wenn ein zugehöriges Sprachelement aufgesucht wird. Ebenso wie beim Visual Browser 12 spezifiziert "cohref" für den Voice Browser 13 die aufzusuchende URL, wenn ein zugehöriger visueller Link oder eine Schaltfläche ausgewählt wird. "Coaction" spezifiziert für den Voice Browser 13 die aufzusuchende URL, wenn ein zugehöriges visuelles Formular abgeschickt wurde. Schließlich spezifiziert "covisit" für den Voice Browser 13 die aufzusuchende URL, wenn ein zugehöriges visuelles Dokument geladen wird oder ein visueller Anker dargestellt wird.

[0040] Es ist anzumerken, dass sowohl die visuellen als auch die Sprachinhalte in einem einzelnen, zusammengesetzten netzbasierten Dokument enthalten sein können. Im Fall eines zusammengesetzten netzbasierten Dokuments können sowohl der Voice Browser 13 als auch der Visual Browser 12 das zusammengesetzte Dokument syntaktisch analysieren und die darin enthaltenen Inhalte darstellen, die für die Darstellung in dem jeweiligen Browser geeignet sind. Darüber hinaus kann sowohl der Voice Browser 13 als auch der Visual Browser 12 in dem zusammengesetzten Dokument die jeweils durch "cohref" und "conext" spezifizierten elektronischen Inhalte identifizieren. Es folgt ein Beispiel eines zusammengesetzten Dokuments:

```

40 <HTML>
    <HEAD><TITLE> Co-Browsing-Demonstration SBU-
    eingebetteter Formatierungszeichen</TITLE></HEAD>
45
    <vxml version="1.0">
50       <form name="Willkommen">
           <block>Willkommen bei der SBU Multi-
           Modus-Browser-Demonstration!
55           <goto next="#1"/>
60
65

```

```

</block>
</form>
<menu id="1">
    <prompt>Sagen Sie zur
    Benutzerauthentifizierung
    "Test".</prompt>
    <grammar type="text/jsgf">
        continue
    </grammar>
    <choice next="mainmenu.vxml"
    conext="mainmenu.html">
        test
    </choice>
</menu>
</vxml>

<BODY LEFTMARGIN="0" TOPMARGIN="0">
    <A HREF="mainmenu.html"
    cohref="mainmenu.vxml">
        <IMG SRC="splashscreen.jpg" WIDTH="314"
        HEIGHT="200"/>
    </A>
</BODY>
</HTML>

```

[0041] Fig. 3 veranschaulicht einen Multi-Modus-Browser 150 zur Implementierung des Verfahrens der Erfindung. Wie aus der Figur ersichtlich ist, kann ein Multi-Modus-Browser 150 gemäß den erfindungsgemäßen Anordnungen drei verschiedene Komponenten umfassen: einen Visual Browser 120, einen Voice Browser 130 und eine Kopplungsschnittstelle 100 zur Kopplung des Voice Browsers 130 mit dem Visual Browser 120 durch Verwalten der Synchronisation zwischen den beiden Browsern. Zusätzlich kann der Multi-Modus-Browser 150 eine Steuerroutine für zeitkritische Ereignisse 140 zur Verarbeitung zeitkritischer Informationen wie etwa Statusinformationen einschließen. Bedeutsamerweise sind die Komponenten Visual Browser 120 und Voice Browser 130 unabhängig. Jede interpretiert ihren eigenen Formatierungssteuerzeichenstrom, um einem Benutzer 160 unabhängige aber zueinander komplementäre Informationen darzustellen. Gemäß dem Verfahren der Erfindung können die Browser miteinander durch die Kopplungsschnittstelle 100 einen Dialog führen.

[0042] In der bevorzugten Ausführungsart des Multi-Modus-Browsers 150 ist der Visual Browser 120 ein Java-basierter Mikro-Browser, der für die Ausführung in einer eingebetteten Umgebung entworfen wurde. Der Mikro-Browser kann mit Hilfe von VisualAge® for Embedded Systems® der IBM Corporation of Armonk, New York, implementiert werden. VisualAge for Embedded Systems ergänzt die Java-Benutzerschnittstellenunterstützung mit der Anzeigetechnologie Degas®. Degas ist insbesondere eine Javabasierte Entwicklungsumgebung, die zum Erstellen GUIeingebetteter Anwendungen zur Ausführung auf dem eingebetteten Betriebssystem QNX Neutrino verwendet wird. Degas kann dazu verwendet werden, Bildschirmanzeigen zu erstellen, die einzigartige grafische Ansichten und ein realistisches 3D-Aussehen und -Verhalten besitzen. Genau gesagt kann ein Entwickler, der das Degas-Gerüst verwendet, visuelle Elemente für eine Ziel-einheit konstruieren, die Klassen, die solche Elemente implementieren, in Unterklassen unterteilen, und einen GUI-Schnittstellen-Prototyp entwerfen, der solche Elemente verwendet.

[0043] Der Mikro-Browser kann die Degas-Klasse Standardapplikationen (DefaultApplication) erweitern und ein HTML-Ansicht-/Controller-Paar als Erweiterung der Degas-Klassen Standardansicht (DefaultView) bzw. Standardcontroller (DefaultController) implementieren. Der HTML-Controller kann Benutzerereignisse von der Anzeige an die Mikro-Browser-Anwendung weitergeben, die wiederum die HTML-Ansicht aktualisieren kann. Eine detailliertere Erörte-

nung des Degas-Gerüsts findet sich in GUI Navigation in Embedded Automotive Systems, von Kern Ann Hamberg und Brian Neal, <http://www.tradespeak.com/html/docs/2030.html> (tradespeak 2000), das durch die Bezugnahme Bestandteil dieses Patents ist.

- [0044] In Fig. 3 ist die Steueroutine für zeitkritische Ereignisse 140 eine weitere Komponente des Multi-Modus-Browsers 150. Die Steueroutine für zeitkritische Ereignisse 140 kann auf einem IP-Port Formatierungssteuerzeichen empfangen, die entweder an den Visual Browser 120, den Voice Browser 130 oder an beide über die Kopplungsschnittstelle 100 weitergeleitet werden können. Das Formatierungssteuerzeichen entspricht zeitkritischen Informationen (hier als Statusinformationen bezeichnet), die die aktuelle Anwendung unterbrechen können. Um zu bestimmen, welcher Browser die Statusinformationen anzeigen soll, umfasst das Formatierungssteuerzeichen vorzugsweise einen Inhalts-Typ, oder MIME-Typ, der sich im Kopf des Dokuments, das das mit dem Formatierungssteuerzeichen enthält, befindet. Gemäß der Konvention folgen zwei leere Zeilen auf den Inhalts-Typ, entweder oberhalb der Anfangskennung des HTML- oder oberhalb des VXML-Dokuments. In der bevorzugten Ausführungsart enthält das Formatierungssteuerzeichen, das Statusinformationen enthält, die von dem Multi-Modus-Browser 150 empfangen werden, vorzugsweise einen Inhalts-Typ VXML, CHTML oder EVXML.
- [0045] Beim Betrieb kann der Benutzer 160, basierend auf einer Benutzerinteraktion mit Sprach- oder visuellen Inhalten, zu einer neuen Adresse navigieren. Beim Übergang eines jeden Browsers von einer URL zu einer anderen kann sich der eine Browser mit dem anderen Browser über die Koordinationsformatierungsattribute, die zu den jeweiligen Formatierungssprachen hinzugefügt sind, beim Durchsuchen koordinieren. Diese Elementattribute können den Browser anweisen, den anderen Browser zu benachrichtigen, wenn eine komplementäre URL ("co-URL") aufgesucht werden soll. Wenn der Benutzer 160 z. B. einen Link im Visual Browser 120 auswählt, sucht der Visual Browser 120 die URL der visuellen Seite auf, die durch den Link spezifiziert wird. Um den Voice Browser 130 anzuweisen, die entsprechende Sprachseite aufzusuchen, kann ein Anwendungsentwickler ein Attribut "covisit" hinzufügen, das die URL der gewünschten Sprachseite zur HTML-Anker oder -Formular-Kennung in der visuellen Seite enthält. Analog sucht der Voice Browser die URL der Sprachseite auf, die durch den Link spezifiziert wird, wenn der Benutzer 160 einen Link im Voice Browser 130 auswählt. Um den Visual Browser 120 anzuweisen, die entsprechende visuelle Seite aufzusuchen, kann der Anwendungsentwickler ein Attribut "comenu", "coform" oder "conext" hinzufügen, das die URL der gewünschten visuellen Seite zur VXML-Kennung in der Sprachseite enthält.
- [0046] Somit kann jede Browserkomponente (Voice und Visual) ihren Formatierungssteuerzeichenstrom interpretieren und ihren eigenen Übergang von einem Dokument oder Dokumentelement (URL) zu einem anderen Dokument oder Dokumentelement machen. Die Dokumentelemente enthalten Formatierungsattribute, die die Übergänge eines jeden Browsers auf den anderen an Punkten, die durch den Anwendungsentwickler gewählt werden, synchronisieren.
- [0047] Somit können die Koordinationsformatierungsattribute, z. B. die Kennungen "cohref" und "conext", zur Kopplung eines Visual Browsers an einen Voice Browser verwendet werden, ohne eine umfassende Änderung einer der beiden zu erfordern. Die Inhaltsformatierungssteuerzeichen selbst können vielmehr als Kopplungsagent agieren. Somit kann, wenn z. B. ein Benutzer auf eine visuelle Anzeige einer Abbildung in einem Visual Browser mit der Maus klickt, eine Kennung "co-target", die zur Abbildung gehört, einem Voice Browser anzeigen, dass er die der Abbildung zugeordneten Audioinhalte laden und abspielen soll. Umgekehrt kann ein Voice Browser einen Benutzer auffordern, den Namen eines geografischen Orts auszusprechen, für den der Voice Browser Zusatzinformationen akustisch zur Verfügung stellen kann. Als Antwort auf den Empfang der Spracheingabe, die einen geografischen Ort bezeichnet, kann eine Kennung "co-target", die zur Sprachaufforderung gehört, einem Visual Browser anzeigen, dass er eine Karte des Orts des angegebenen geografischen Orts laden und anzeigen soll.

Patentansprüche

1. Verfahren zum gleichzeitigen Zugreifen auf netzbasierte elektronische Inhalte in einem Voice Browser und einem Visual Browser, das folgende Schritte umfasst:
Abrufen eines netzbasierten Dokuments, das für die Anzeige im Visual Browser formatiert ist;
Identifizieren eines Bezugs auf den Voice Browser in dem abgerufenen Dokument, wobei der Bezug elektronische Inhalte, die für die akustische Darstellung im Voice Browser formatiert sind, spezifiziert; und
übermitteln des Bezugs an den Voice Browser;
Abrufen der spezifizierten elektronischen Inhalte durch den Voice Browser und akustische Darstellung der elektronischen Inhalte im Voice Browser;
visuelle Darstellung des netzbasierten Dokuments durch den Visual Browser gleichzeitig mit der akustischen Darstellung.
2. Verfahren gemäß Anspruch 1, bei dem das netzbasierte Dokument, das für die visuelle Darstellung im Visual Browser formatiert ist, ein HTML-formatiertes Dokument ist und der Visual Browser ein Web-Browser ist.
3. Verfahren gemäß Anspruch 1, bei dem die elektronischen Inhalte, die für die akustische Darstellung im Voice Browser formatiert sind, VoiceXML-formatierte elektronische Inhalte sind.
4. Verfahren gemäß Anspruch 2, bei dem der Bezug ein Koordinationsformatierungsattribut ist.
5. Verfahren gemäß Anspruch 1, bei dem der Bezug ein netzbasiertes Dokument spezifiziert, das die elektronischen Inhalte enthält, die für die akustische Darstellung im Voice Browser formatiert sind, wobei der Voice Browser das spezifizierte netzbasierte Dokument über das Computernetzwerk abrufen und die elektronischen Inhalte, die in dem spezifizierten netzbasierten Dokument enthalten sind, akustisch darstellen kann, während der Visual Browser das netzbasierte Dokument, das für die visuelle Darstellung formatiert ist, visuell darstellen kann.
6. Verfahren zum gleichzeitigen Zugreifen auf netzbasierte elektronische Inhalte in einem Voice Browser und einem Visual Browser, das Folgendes umfasst:
Abrufen eines netzbasierten elektronischen Dokuments, das für akustische Darstellung im Voice Browser formatiert ist;

- Identifizieren eines Bezugs auf den Visual Browser in dem abgerufenen Dokument, wobei der Bezug die elektronischen Inhalte, die für die visuelle Darstellung im Visual Browser formatiert sind, spezifiziert; und Übermitteln des Bezugs an den Visual Browser;
- Abrufen der spezifizierten elektronischen Inhalte durch den Visual Browser und visuelle Darstellung der elektronischen Inhalte im Visual Browser;
- akustische Darstellung des netzbasierten Dokuments durch den Voice Browser gleichzeitig mit der visuellen Darstellung.
7. Verfahren gemäß Anspruch 6, bei dem das netzbasierte Dokument, das für die akustische Darstellung im Voice Browser formatiert ist, ein VoiceXML-Dokument ist.
8. Verfahren gemäß Anspruch 6, bei dem die elektronischen Inhalte, die für die visuelle Darstellung im Visual Browser formatiert sind, HTML-formatierte elektronische Inhalte sind und der Visual Browser ein Web-Browser ist.
9. Verfahren gemäß Anspruch 6, bei dem der Bezug ein netzbasiertes Dokument spezifiziert, das die elektronischen Inhalte enthält, die für die visuelle Darstellung im Visual Browser formatiert sind, wobei der Visual Browser das spezifizierte netzbasierte Dokument über das Computerkommunikationsnetz abrufen und die elektronischen Inhalte, die in dem netzbasierten Dokument enthalten sind, visuell darstellen kann, während der Voice Browser das netzbasierte Dokument, das für die akustische Darstellung formatiert ist, akustisch darstellen kann.
10. Maschinenlesbarer Speicher, auf dem ein Computerprogramm zum gleichzeitigen Zugreifen auf netzbasierte elektronische Inhalte in einem Visual Browser und einem Voice Browser gespeichert ist, wobei das Computerprogramm eine Vielzahl von Codeabschnitten besitzt, die durch eine Maschine ausführbar sind, um die Maschine zur Durchführung der folgenden Schritte zu veranlassen:
- Abrufen eines netzbasierten Dokuments, das für die Anzeige im Visual Browser formatiert ist;
- Identifizieren eines Bezugs auf den Voice Browser in dem abgerufenen Dokument, wobei der Bezug elektronische Inhalte, die für die akustische Darstellung im Voice Browser formatiert sind, spezifiziert; und
- Übermitteln des Bezugs an den Voice Browser;
- Abrufen der spezifizierten elektronischen Inhalte durch den Voice Browser und akustische Darstellung der elektronischen Inhalte im Voice Browser;
- visuelle Darstellung des netzbasierten Dokuments durch den Visual Browser gleichzeitig mit der akustischen Darstellung.
11. Maschinenlesbarer Speicher gemäß Anspruch 10, bei dem das netzbasierte Dokument, das für die visuelle Darstellung im Visual Browser formatiert ist, ein HTML-Dokument ist und der Visual Browser ein Web-Browser ist.
12. Maschinenlesbarer Speicher gemäß Anspruch 10, in dem die elektronischen Inhalte, die für die akustische Darstellung im Voice Browser formatiert sind, VoiceXML-formatierte elektronische Inhalte sind.
13. Maschinenlesbarer Speicher gemäß Anspruch 11, bei dem der Bezug ein Koordinationsformatierungsattribut ist.
14. Maschinenlesbarer Speicher gemäß Anspruch 10, bei dem der Bezug ein netzbasiertes Dokument spezifiziert, das die elektronischen Inhalte enthält, die für die akustische Darstellung im Voice Browser formatiert sind, wobei der Voice Browser das spezifizierte netzbasierte Dokument über das Computerkommunikationsnetz abrufen und die elektronischen Inhalte, die in dem spezifizierten netzbasierten Dokument enthalten sind, akustisch darstellen kann, während der Visual Browser das netzbasierte Dokument, das für die visuelle Darstellung formatiert ist, visuell darstellen kann.
15. Maschinenlesbarer Speicher, auf dem ein Computerprogramm zum gleichzeitigen Zugreifen auf netzbasierte elektronische Inhalte in einem Visual Browser und einem Voice Browser gespeichert ist, wobei das Computerprogramm eine Vielzahl von Codeabschnitten, die durch eine Maschine ausführbar sind, besitzt, um die Maschine zur Durchführung der folgenden Schritte zu veranlassen:
- Abrufen eines netzbasierten Dokuments, das für die akustische Darstellung im Voice Browser formatiert ist;
- Identifizieren eines Bezugs auf den Visual Browser in dem abgerufenen Dokument, wobei der Bezug die elektronischen Inhalte, die für die visuelle Darstellung im Visual Browser formatiert sind, spezifiziert;
- Übermitteln des Bezugs an den Visual Browser;
- Abrufen der spezifizierten elektronischen Inhalte durch den Visual Browser und visuelle Darstellung der elektronischen Inhalte im Visual Browser;
- akustische Darstellung des netzbasierten Dokuments durch den Voice Browser gleichzeitig mit der visuellen Darstellung.
16. Maschinenlesbarer Speicher gemäß Anspruch 15, bei dem das netzbasierte Dokument, das für die akustische Darstellung im Voice Browser formatiert ist, ein VoiceXML-formatiertes Dokument ist.
17. Maschinenlesbarer Speicher gemäß Anspruch 15, bei dem die elektronischen Inhalte, die für die visuelle Darstellung im Visual Browser formatiert sind, HTML-formatierte elektronische Inhalte sind und der Visual Browser ein Web-Browser ist.
18. Maschinenlesbarer Speicher gemäß Anspruch 15, bei dem der Bezug ein netzbasiertes Dokument spezifiziert, das die elektronischen Inhalte enthält, die für die visuelle Darstellung im Visual Browser formatiert sind, wodurch der Visual Browser das spezifizierte netzbasierte Dokument über das Computerkommunikationsnetz abrufen und die elektronischen Inhalte, die in dem spezifizierten netzbasierten Dokument enthalten sind, visuell darstellen kann, während der Voice Browser das netzbasierte Dokument, das für die akustische Darstellung formatiert ist, akustisch darstellen kann.
19. Ein Multi-Modus-Browser, der Folgendes umfasst:
- einen Visual Browser zur visuellen Darstellung visueller Inhalte, die von einem Computerkommunikationsnetz abgerufen wurden;
- einen Voice Browser zur akustischen Darstellung von Sprachinhalten, die von dem Computerkommunikationsnetz

abgerufen wurden; und

eine Kopplungsschnittstelle zur Synchronisation der visuellen und akustischen Darstellung der visuellen und der Sprachinhalte in jedem der jeweiligen Visual bzw. Voice Browser.

20. Multi-Modus-Browser gemäß Anspruch 19, bei dem die visuellen Inhalte HTML-formatierte Inhalte und die Sprachinhalte VoiceXML-formatierte Inhalte sind.

21. Multi-Modus-Browser gemäß Anspruch 19, der des Weiteren umfasst:

eine Steuerroutine für zeitkritische Ereignisse,

wobei die Steuerroutine für zeitkritische Ereignisse zeitkritische Inhalte empfängt,

die Steuerroutine für zeitkritische Ereignisse in den zeitkritischen Inhalten einen Inhalts-Typ-Indikator identifiziert,

die Steuerroutine für zeitkritische Ereignisse die zeitkritischen Inhalte an den Visual Browser weiterleitet, wenn der Inhalts-Typ-Indikator visuelle Inhalte anzeigt,

die Steuerroutine für zeitkritische Ereignisse die zeitkritischen Inhalte an den Voice Browser weiterleitet, wenn der Inhalts-Typ-Indikator Sprachinhalte anzeigt.

22. Verfahren zum Ändern eines netzbasierten Dokuments zur Unterstützung des gleichzeitigen Zugriffs auf netzbasierte Sprach- und visuelle Inhalte in einem Voice Browser und einem Visual Browser, das die folgenden Schritte umfasst:

Integrieren von visuellen Inhalten in ein netzbasiertes Dokument;

Formatieren des netzbasierten Dokuments für die visuelle Darstellung in einem Visual Browser; und

Einfügen von mindestens einer Formatierungskennung in das netzbasierte Dokument, wobei die Formatierungskennung ein Koordinationsformatierungsattribut enthält, das ein netzbasiertes Dokument spezifiziert, das für die akustische Darstellung in einem Voice Browser formatiert ist,

wodurch der Visual Browser bei der Wiedergabe des netzbasierten Dokuments, das für die visuelle Darstellung formatiert ist, das Koordinationsformatierungsattribut in der Formatierungskennung identifizieren und einen Bezug auf das spezifizierte netzbasierte Dokument an den Voice Browser übertragen kann, was den Voice Browser veranlasst, das spezifizierte netzbasierte Dokument abzurufen und es gleichzeitig mit der visuellen Darstellung des netzbasierten Dokuments, das für die visuelle Darstellung in dem Visual Browser formatiert ist, akustisch darzustellen.

23. Verfahren gemäß Anspruch 22, bei dem das netzbasierte Dokument, das für die visuelle Darstellung in einem Visual Browser formatiert ist, ein HTML-formatiertes Dokument ist, und der Visual Browser ein Web-Browser ist.

24. Verfahren gemäß Anspruch 22, bei dem das netzbasierte Dokument, das für die akustische Darstellung in einem Voice Browser formatiert ist, ein VoiceXML-formatiertes Dokument ist und der Voice Browser ein VoiceXML-Browser ist.

25. Verfahren zum Ändern eines netzbasierten Dokuments zur Unterstützung des gleichzeitigen Zugriffs auf netzbasierte Sprach- und visuelle Inhalte in einem Voice Browser und einem Visual Browser, das die folgenden Schritte umfasst:

Integrieren von Sprachinhalten in ein netzbasiertes Dokument;

Formatieren des netzbasierten Dokuments für die akustische Darstellung in einem Voice Browser;

Einfügen von mindestens einer Formatierungskennung in das netzbasierte Dokument;

wobei die Formatierungskennung ein Koordinationsformatierungsattribut enthält, das das netzbasierte Dokument spezifiziert, das für die visuelle Darstellung in einem Visual Browser formatiert ist,

wodurch der Voice Browser bei der Wiedergabe des netzbasierten Dokuments, das für die akustische Darstellung formatiert ist, das Koordinationsformatierungsattribut in der Formatierungskennung identifizieren und einen Bezug auf das spezifizierte netzbasierte Dokument an den Visual Browser übertragen kann, was den Visual Browser veranlasst, das spezifizierte netzbasierte Dokument abzurufen und das spezifizierte netzbasierte Dokument gleichzeitig mit der akustischen Darstellung des netzbasierten Dokuments, das für die akustische Darstellung in dem Voice Browser formatiert ist, visuell darzustellen.

26. Verfahren gemäß Anspruch 25, bei dem das netzbasierte Dokument, das für die visuelle Darstellung in einem Visual Browser formatiert ist, ein HTML-formatiertes Dokument ist und der Visual Browser ein Web-Browser ist.

27. Verfahren gemäß Anspruch 25, bei dem das netzbasierte Dokument, das für die akustische Darstellung in einem Voice Browser formatiert ist, ein VoiceXML-formatiertes Dokument ist und der Voice Browser ein VoiceXML-Browser ist.

Hierzu 3 Seite(n) Zeichnungen

- Leerseite -

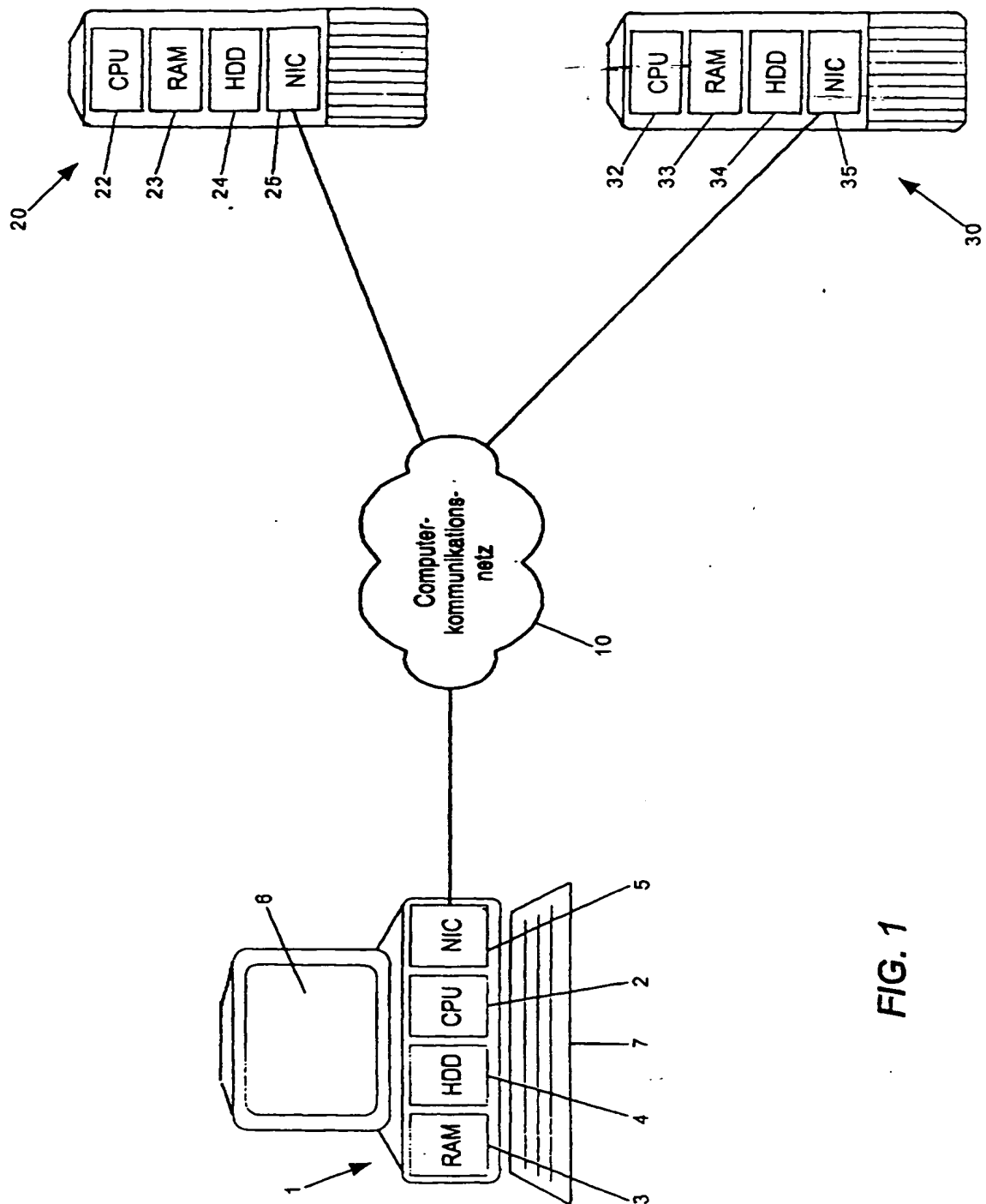


FIG. 1

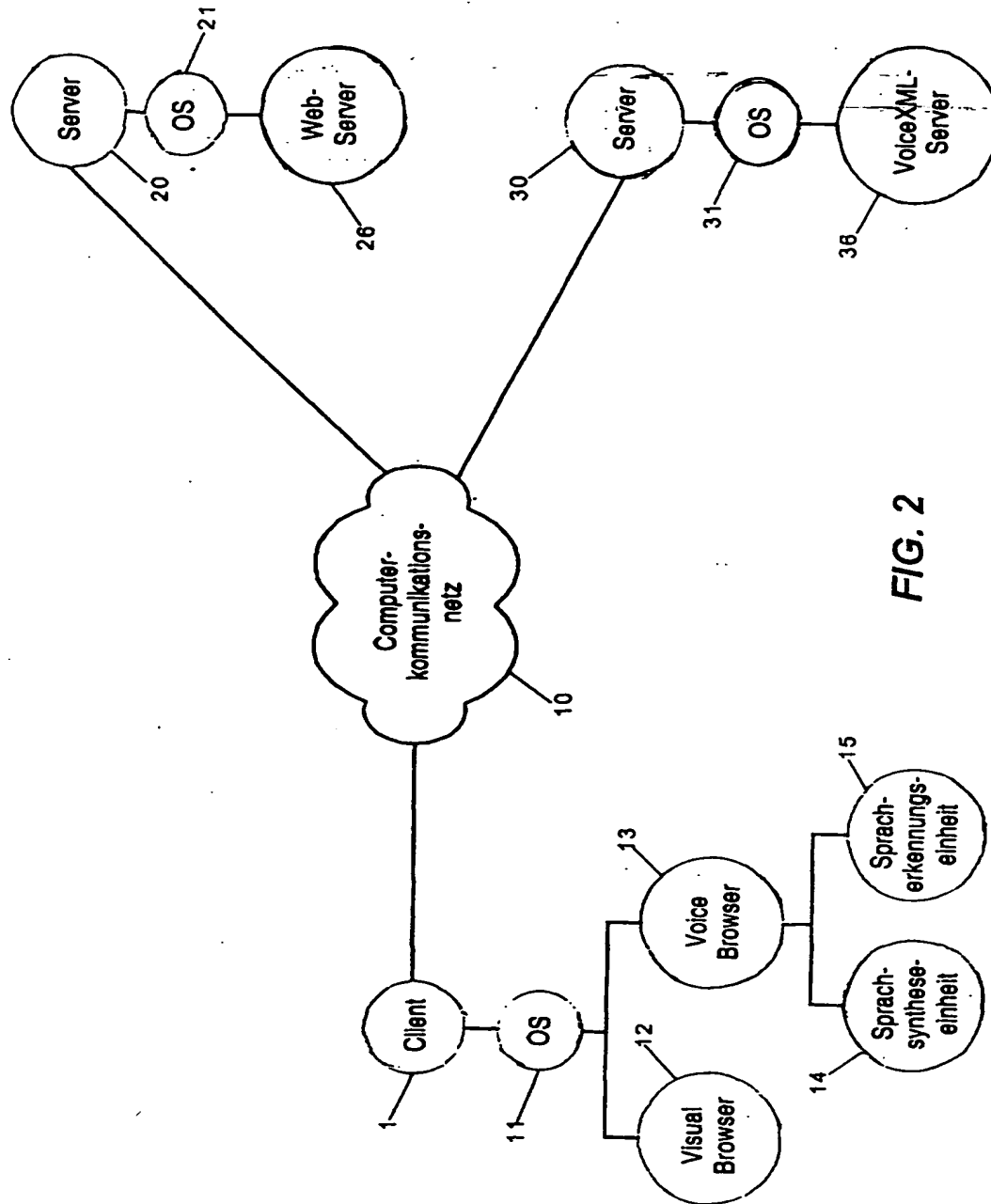


FIG. 2

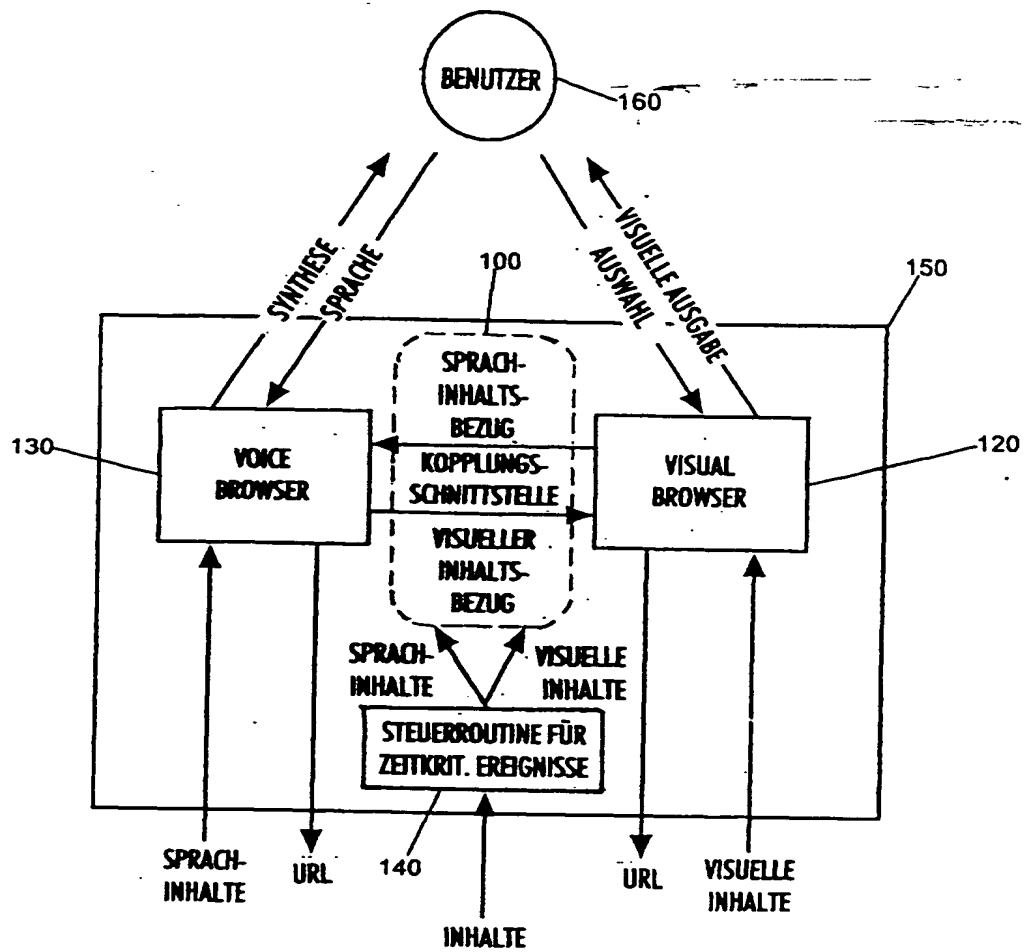


FIG. 3